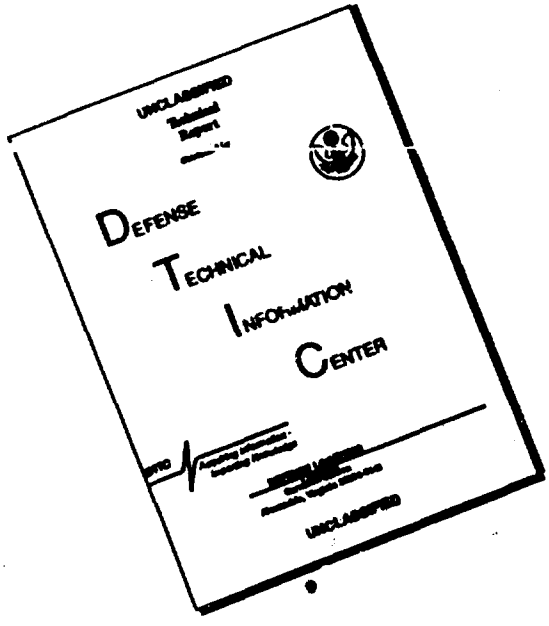
TDD - TR-10-001

SURVEY OF RESEARCH ON READABILITY OF TECHNICAL PUBLICATIONS

H. R. BOOHER

APRIL 1971

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Naval Air Systems Command	Unclassified
	2b. GROUP

REPORT TITLE

Survey of Research on Readability of Technical Publications

DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final Report - Jan 1970 - Jan 1971

AUTHOR(S) (First name, middle initial, last name)

Harold R. Booher

REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1971	11	21

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	TDD-TR-10-001
c. In-house	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED	Naval Air Systems Command

13. ABSTRACT

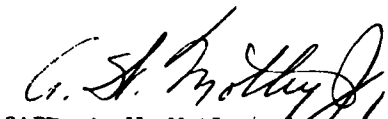
Since the utility of technical manuals is greatly influenced by their reading difficulty, a survey of factors affecting readability of technical publications was conducted. The work done by George Klare was relied on to a large extent for a bibliography of methods of measuring readability prior to 1966. The survey covers additional research done since that time with particular applications toward reading technical materials. The report concludes with suggested areas for Human Factors research in readability of technical publications.

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Reading						
Readability						
Technical Manuals						
Technical Publications						
Writing						
Technical Writing						
Reading Education						
Education						
Training						
Human Factors						
Readability Measurements						

FOREWORD

This study represents a portion of the Human Factor research and development program of the Technical Documentation Department under AIRTASK A415 415D-2234-1415000000. As part of a research effort to predict the utility of technical manuals in NAVAIR, a survey of the literature on existing readability techniques was conducted. The report covers research performed between January 1970 and January 1971.

This technical report has been reviewed and is approved.



CAPT. A. W. Motley, Jr.
Director
Technical Documentation Division
Naval Air Systems Command

ABSTRACT

Since the utility of technical manuals is greatly influenced by their reading difficulty, a survey of factors affecting readability of technical publications was conducted. The work done by George Klare was relied on to a large extent for a bibliography of methods of measuring readability prior to 1966. The survey covers additional research done since that time with particular applications toward reading technical materials. The report concludes with suggested areas for Human Factors research in readability of technical publications.

Survey of Research on Readability of Technical Publications

Introduction

The literature on readability research is vast but only a very small portion has any direct relation to technical publications. The research on readability has for the most part been designed around isolating factors affecting reading efficiency, judgements of difficulty, levels of comprehension, learning and retention. The scope of this report is limited to major studies in the readability literature which appears to bear most directly on problems associated with the use of technical publications in personnel subsystems.

In personnel subsystems a technical publication is often defined as any device for storing information. It can be a paper manual, a microform reader-printer, a visual-audio instructional system or a computer storage bank with a printer or display readout. The technical publication is used for one or more of three major purposes: for training, as a reference in supplementing the human memory, or for cooperative aiding (e.g. checklists, step-by-step procedures).

There are three general categories into which most readability studies fall. First of all there is a sizable group of studies which primarily indicate the legibility of either hand writing or typography. These are almost entirely perceptual in nature. Human Factors studies on readability almost always refer to the legibility of alphanumerics on visual displays. Another area of research on readability is ease of reading due either to interest value or pleasantness of writing. Only one study was found in this area with technical material as content. (Klare, Mabry, Gustafson, 1955). The remaining category of literature on readability covers work done to indicate ease of understanding due to the style of writing. Findings from this last category are of most interest from a technical publication view point.

This report covers therefore, those classical readability factors which apply to reading material having technical content. In addition it provides a summary of major techniques which can be used in the evaluation of technical publication readability. The report concludes with a discussion of Human Factors research needed in readability of technical publications.

Classical Readability Factors

According to Klare (1963) the first measures of readability recorded were that of the Talmudists in 900 A.D. They made use of word and idea

counts so that the frequency of occurance could be used to distinguish usual from unusual meanings. In the U. S., educators have been interested in readability for sometime - particularly with regard to children's readers. Word counts were popular as early as 1840 in determining reading levels of the McGuffey readers.

Klare cites 1921 as a milestone in readability. During that year Thorndike published "The Teachers Wordbook" wherein he tabulated the frequency with which words occur in print. This probably more than any other one thing influenced the teaching of vocabulary in schools and provided the base for readability formulas development.

The usual factors in reading material which have been idenified as important criteria for reading difficulty are word factors and sentence factors. Apparently word and sentence factors explain about as much of the variance in performance on reading efficiency as can be explained with additional style factors. According to Klare (1963) the most accurate formula in predicting grade levels is the Dale - Chall formula which uses only two factors - word difficulty and sentence length. In designing an automated readability index Smith and Senter (1967) provided a way of mechanically tabulating these two factors for passages as they are typed on a standard typewriter.

Word frequency correlates most highly with recognition time. Soloman and Postman (1952, in Klare, 1963) using artificial words (Turkish) not previously seen by the subject, built up word frequency usage by controlling the number of tachistoscopic exposures to the words. They found that recognition thresholds vary inversely with frequency of prior usage. It was not determined, however, to what extent frequency of prior exposure and frequency of prior response interacted in the relationship.

The older or more frequently words are used, the shorter they tend to become. Television has become TV, radio detecting and ranging is radar and car is all that is left of the horseless carriage. Also it is generally found that the shorter the word, the shorter the recognition time. More often than not, therefore, shorter words are judged easier than longer words. (Klare, 1963).

Word frequency and familiarity of words to a reader have a close relationship, but it is not a perfectly linear one. Klare (1963) notes that there is a law of diminishing returns at work here. Beyond some point of high frequency one additional exposure will have little effect upon recognition, whereas at low frequencies one additional exposure can be highly effective in producing faster recognition times.

The length of the sentence is the next most important factor in the style of writing for reading efficiency. This is partly due to structure of sentences, (complicated sentences are generally longer than simple sentences) but not entirely. The length of the sentence itself because of perceptual memory span makes sentence length a major factor in readability formulas regardless of the sentence structure. Miller and Selfridge, (1950, in Klare, 1963) showed that memory span is closely related to the extent to which organization of verbal context approximates English. The percent of recall increased as the order of approximation increased (Nonsense to English). Recall decreased, however, as sentence length increased.

A sentence factor which has been studied in several different ways is redundancy. Redundancy is the extent to which a given unit of language is determined by nearby units. (Klare, 1963). It has been approached experimentally by the degree to which the words approximate English (Miller & Selfridge, 1950), by deleting letters from words (Chapanis, 1954), or by deleting words (Taylor, 1953; Bormuth, 1967; Federman, MacPherson, and Siegel, 1970). The latter method; known as the "Cloze" technique, has been shown to be an effective method of evaluating surface and submarine personnel comprehension of Sonar manuals. When letters are deleted from words, passages that are easy to reconstruct are not necessarily those easiest to read. Chapanis (1954) gave 13 English prose passages to subjects using both random and regular patterns. He found that in general people do poorly in reconstructing passages when the amount of material deleted is 30% or more.

Special consideration of the reader is required in assessing the readability of written material. Reading level roughly corresponds to the last school grade completed. This will vary depending on any special experience the reader has with the material and with general vocabulary and IQ ratings. (Klare, 1963). Both high-ability and low-ability readers can benefit from more readable material, however. No matter what a reader's skill level, if the same information content is presented in a more readable manner, he will be able to read it with greater ease (Klare, 1963).

The difference between easy and hard versions of material in terms of test scores, may not be reflected in comprehension tests. This is particularly true with technical material. A technician's special knowledge of electronics may allow him to get the same comprehension score on electronics passages regardless of style difficulty. (Klare, Mabry, Gustafson, 1955). When an eye-movement camera is used, however, easy style technical material produced higher scores than the hard style on words per fixation (Klare, Shuford, Nichols, 1957).

Motivation is a major factor in comprehension, learning or retaining information. Subjects with either strong or weak sets to learn material will read faster and with fewer fixations only on easy versions. Reading speed may decrease on hard versions, with increased learning motivation. (Klare, 1963). But Klare (1963) indicates that recall scores will be relatively higher on both hard and easy versions where there is motivation to learn.

A principle of least effort is also at work with most readers. All other things being equal, a reader under low motivation will choose material around two grade levels below his level when reading for pleasure (Klare, 1963). In other words, technical readers are not likely to read difficult technical information (e.g. circuit theory) unless the topic is very interesting, very necessary to something they value highly (e.g. health, life, family) or they are strongly motivated to learn.

Most of the readability research has been primarily on the difficulty of style of writing. Klare (1963) discusses other important factors present in determining the readability of a publication. Broadly these are content, format, and organization of information presented.

In summary, the following general rules can be applied to technical publications.

1. In considering reading speed and efficiency, more readable material provides consistantly for an increase. This holds true for most educational levels and amounts of background readers have. This is particularly important in motivating voluntary reading of difficult material such as the theory of operation found in maintenance manuals.

2. Readers as a group tend to rank materials in terms of readability ease in much the same order as a writer using readability formulas. Readability formulas can give gross indications of acceptability of material.

3. If a publication is to attract a wide readership (e.g. operators - maintainers of equipment) there will be low education levels involved. There will be generally a low set to learn and a large amount of the principle of least effort.

4. Readability formulas will not generally predict how well material will be comprehended. No matter what the style, the material may be comprehended the same because a reader may be highly motivated to learn material, he may have a large amount of experience or background with the topic being covered, or he may read a passage over and over until he understands what is meant. For low motivated, inexperienced, and time limited individuals, however, the more readable the style, the better the comprehension, rate of learning, and level of retention.

Readability/Comprehensibility Measurement Techniques

1. Prediction formulas - The evaluation of readability or comprehensibility of textual material may be determined with one or more of the following methods: expert judgement, simulated field conditions, prediction formulas and the "cloze" technique.

In the past prediction formulas have been used most widely in assessing educational materials. Expert judgements or simulated field conditions are usually employed however, in verifying technical publications. This is partly because readability formulas appear to be weak as indicators of readability of technical material (Ross, 1959) and non-fiction specialized material like physics text books (Marshall, 1957).

Probably the most popular formula is that of Rudolf Flesch (1948). It has attained great popularity because it is a simple formula consisting of only 3 factors; it correlates highly (.74) with McCall-Crabbs Standard Test Lessons; and Flesch popularized it with a series of articles and books getting support of writers in journalism, government and business. The general procedure is:

(1) Select samples of 100 words throughout material.
(2) Compute average sentence length in words (Xs)
(3) Count affixes Xm
(4) Count personal reference (Xh)

The original formula (1943) is:

$$X = .07\ Xm + .07\ Xs - .05\ Xh + 3.27$$

There is a shortened formula for reading ease $RE = 206.835 - .846\ wl - 1.015\ sl$

Where:

wl - number of syllables per 100 words
sl - avg. number of words per sentence

The shortened formula probably is the most applicable to technical publications because it eliminates human interest factor. In technical writing human interest (i.e. you turn the crank) versions have been judged less acceptable by airmen. (Klare, Mabry, Gustafson, May 1955)

The next most popular formula is the Dale-Chall (1948).

The Dale-Chall formula is:

$$X_{c_{50}} = .1579\ X_1 + .0496X_2 + 3.6365$$

Where:

X_2 = average sentence length in words
$X_{c_{50}}$ = Reading grade score of pupil who could answer one-half test questions on McCall - Crabb Test

X_1 = % of words outside Dale list of 3000

A formula which has received acceptance in Air Force Standard Writing guides is the Fog index discussed in Grunning (1952). The procedure is similar but simpler than the Flesch method.

Fog Index

(1) Take 100 word samples
(2) Divide the number of words by number of sentences to get average sentence length
(3) Count number of words by three or more syllables
(4) Total 2 factors and multiply by .4

A final formula is the Gray & Leary

$$X_1 = -.01029X2 + .009012X_5 - .02094X6 - .03313X7 - .01485X_8 + 3.774$$

The work of Gray & Leary (1935) in Klare, 1963 is considered a landmark on readability because of the detailed analysis they used.

The above formula is a regression equation developed from a 44 Factor analysis where:

X_1 = Average Comprehension score (Fiction - Non-fiction) for adults of limited reading ability
X_2 = # of difficult hard words not on the Dale list of 769 words
X_5 = # of personal pronouns
X_6 = Average # of words per sentence
X_7 = % of different words
X_8 = # of prepositional phrases

2. Cloze Techniques

A method which is considered a possible break through in techniques

for evaluating readability of technical publication is the "cloze" technique. The procedure was developed by Taylor (1953) but has only recently been applied to the evaluation of technical manuals (Federman, MacPherson and Siegel, 1970). The technique simply requires deletion of words from a passage and then scoring the reader on the percentage of words he can correctly replace. According to Taylor (1957, in Federman, MacPherson and Siegel, 1970) the cloze procedure assumes that there is a high correlation between how readable a piece of writing is and how well it can be understood with words left out. The more a person understands a piece of writing the more likely he can guess the missing words.

The cloze procedure correlates highly with test comprehension scores. Bormuth (1967) for example, reports a product moment correlation of .946 for 4th and 5th grade comprehension on multiple choice test with cloze scores. Bormuth's technique with cloze scores involves selectively removing words from the sample to reduce the effect from redundancy rather than understandability suggesting the missing words. (Klare, 1966).

3. Use Tests, Expert Judgements and Comprehension Tests

Actual use, expert judgement, and comprehension tests have been utilized extensively in evaluating technical manuals (Federman, MacPherson and Siegel). Ross (1959) developed a weighted check list for evaluating technical manuals covering readability of technical manuals. The check list depends upon a combination of expert judgements, the Fog Index, use of peak stress (underlining important words in sentences) and use of personal pronouns in the publication. The checklist also includes factors for layout, organization, decision making procedures and pictoral or schematic illustrations in determining the comprehensibility of technical material. Many factors other than written language enter into the determination of the comprehensibility of technical manuals. These are covered elsewhere. (Folley and Munger, 1961)

There are problems associated with most expert judgements, actual use tests in simulated field conditions, and comprehension tests that make them undesirable as evaluation techniques. Simulating the actual use of a technical manual is costly, inappropriate early in system development, and confounds technician ability with manual inadequacy. Comprehension tests are unstandardized from manual to manual and may not reflect readability faults. Expert judgements are undesirable because of problems in selecting experts, defining valid criteria measures, and in generalizing from expert opinion to the user population. (Federman, MacPherson and Siegel, 1970).

Areas for Research in Readability of Technical Publications

Far too little research has been done with readability of technical publications. As Klare (1963) points out, most readability research is

centered around difficulty of style with almost no research on measures of content. Unless content is known not to change when style is changed, it is difficult to attribute changes in comprehension scores to changes in style.

Chapanis (1965) feels that all of the present readability formulas including the cloze techniques are virtually useless to our problems in most instructional material. He states that we should be interested in the intelligibility or understandability of what is said rather than reading speed, comprehension of text material or judgement of difficulty. The close technique although useful in evaluating technical manual sections appears to be of little value in short instructions. The cloze procedure as presently designed requires large samples (usually 250 words) whereas our samples of language in instructional material are often short.

In experiments which have been designed to find out the best ways of presenting instructions to bring about the specific human actions intended, the results have been gratifying. Only two such experiments have been found, however. Conrad (1962) found that a significantly greater number of phone calls could be successfully transmitted with simplified, clear instructions than could be made with mechanical changes in typography.

Haney (1969) found that when information is action sequenced testing performance is increased. A tabular format with well defined step by step instructions was superior to standard format. Haney emphasized that the information had to be action sequenced. It is not enough to draw lines down a page or put sentences in a column.

Recent research in psycholinguistics such as syntactic language structuring around a "kernel" of language provides a method for developing rules for structuring instructions (Miller, 1962, 1969). Miller (1962) has shown that simple active form sentences are interpreted faster than negative or passive forms. Chapanis (1965) encourages experimental work be done with complicated sentences and then be tested out in terms of the effects on the behavior of ordinary people. Perhaps the degree of complication of a sentence can be quantified in terms of some unit similar to information bits. As sentences have more than one adjective modifying a noun or more than one object of a verb, the number of alternatives increase, therefore decreasing the rate at which the information can be processed.

Chapanis (1965) suggested several areas for human factors research which still need to be carried out. As a measure of intelligibility the cloze technique should be validated on various technical manuals with different manual users. Comparison of the Taylor and Bormuth forms of cloze on technical publications may be valuable in obtaining a more valid evaluation tool. Short instructions built up on the basis of task analysis similar to that of Haney or Conrad should be investigated further.

Criteria for evaluating manuals in terms of the content and understandability of the material is needed. Methods of testing manuals and guides for preparing them are in great demand.

Other research is needed to decide what mixture of pictures, schematics and words provide the most understandable combination.

Special word lists are needed for special purposes. Words that are likely to be within the comprehension of the average maintenance technician should be compiled.

To date there is almost nothing on the readability of foreign languages. As we start designing machines and instruction manuals for use by diverse nationalities, research problems become evident. Perhaps instructions consisting entirely of pictures would be appropriate for guiding specific human actions to be taken in their proper order.

Finally research on what motivates people to use a technical publication is badly needed. How can publications be designed to instill interest, or provide a strong set to learn? Are colors and special bindings motivating? Will a maintenance man prefer comic characteristics to plain pictures? Is a sense of accomplishment with a technical publication rewarding? Will a maintenance man operate buttons on a machine to find information more often than pick up an index to a paper publication? Can special training techniques be designed to encourage motivation to learn how to use technical publications properly?

Systematic research into any of these problems should reveal basic behavioral relationships in a very important but neglected area of man-machine communications - an area which in the past has relied almost exclusively on the individual judgments of technical writers and graphic artists.

References

1. Bormuth, John R. — Comparable Cloze and Multiple Choice Comprehension Test Scores - Journal of Reading 1967, 10 (5), 291-299.

2. Chapanis, Alphonse — The Reconstruction of Abbreviated Printed Messages, Journal of Experimental Psych. 48: 496-510 Dec 1954.

3. Chapanis, Alphonse — Words, Words, Words - Human Factors, 1-17 Feb 1965 AD 632 858.

4. Conrad, R. — The design of information. Occupational Psychology 36, 159-162, 1962.

5. Dale, E. and Chall J.S. — "A Formula for Predicting Readability" Educational Research Bulletin 27: 11-20 January 21, 1948.

6. Flesch, R. F. — A New Readability Yardstick Journal of Applied Psychology 32: 221-33 June 1948.

7. Grunning R. — The Technique of Clear Writing. N.Y. McGraw Hill 1952.

8. Federman, Phillip J. — MacPherson, D. H. Siegel, A. I. Normative Development for Submarine Sonar Manuals; Applied Psychological Services, Science Center, Wayne, Pa. Contract N00014-67-C-0450 NR 196-073/1-2-67 January 1970

9. Folley, John D. Jr. Munger, Sara J. — A Review of the Literature on Design of Informational Job Performance Aids ASD TR-61-549

10. Haney, Richard — The Effect of Format on Functional Testing Performances Human Factors 69 11 (2) 181-188.

11. Klare, George R. — The Measurement of Readability, Iowa State University Press, 1963.

12. Klare, George R. — Comments on Bormuths Readability - Reading Research Quarterly 1966 1(4) 119-125.

13. Klare, G. R. Mabry, J. R. Gustafson, L. M. — "The Relationship of Human Interest to Immediate Retention and to Acceptability of Technical Material, Journal of Applied Psychology 39: 92-95 1955

14. Klare, G. R. Mabry, J. R. Gustafson, L. M. — The relationship of Style Difficulty to Immediate Retention and to Acceptability of Technical Material. Journal Educational Psychology 46: 287-95. May, 1955.

15. Klare, G. R.
Shuford, E. H.
Nichols, W. H.
The relationship of Style Difficulty, Practice, and Ability to Efficiency of Reading and to Retention. Journal Applied Psychology 41: 222-26 1957.

16. Marshall, J. S.
The Relationship Between Readability and Comprehension of High School Physics Text Books. Dissertation Abstracts 17:64 Jan 1957.

17. Miller, George A.
Some Psychological Studies of Grammar. American Psychologist 17, 748-762, 1962.

18. Miller, G. A.
A Psychological Method to Investigate Verbal Concepts. Journal Mathematical Psychology 6(2) 169-191 1969.

19. Ross, Donald A.
Comprehensibility Evaluation of Technical Manuals WADC TN 59-442 July 1969 AD 228 235.

20. Smith, E. A.
Senter, R. J.
Automated Readability Index, Aerospace Medical Research Laboratories, Air Force Systems Command, Wright-Patterson AFB, Ohio. AMRL-TR-66-220 November 1967.

21. Taylor, W. L.
Cloze Procedures: A New tool for Measuring Readability Journalism Quarterly 30. 415-33 Fall 1953.